GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 20, GB4012, doi:10.1029/2005GB002610, 2006

Analysis of ecosystem controls on soil carbon source-sink relationships in the northwest Great Plains

Zhengxi Tan,[1,2] Shuguang Liu,[3] Carol A. Johnston,[1] Jinxun Liu,[3] and Larry L. Tieszen[4]

Received 29 August 2005; revised 11 April 2006; accepted 14 July 2006; published 22 November 2006.

[1] Our ability to forecast the role of ecosystem processes in mitigating global greenhouse effects relies on understanding the driving forces on terrestrial C dynamics. This study evaluated the controls on soil organic C (SOC) changes from 1973 to 2000 in the northwest Great Plains. SOC source-sink relationships were quantified using the General Ensemble Biogeochemical Modeling System (GEMS) based on 40 randomly located 10×10 km^2 sample blocks. These sample blocks were aggregated into cropland, grassland, and forestland groups based on land cover composition within each sample block. Canonical correlation analysis indicated that SOC source-sink relationship from 1973 to 2000 was significantly related to the land cover type while the change rates mainly depended on the baseline SOC level and annual precipitation. Of all selected driving factors, the baseline SOC and nitrogen levels controlled the SOC change rates for the forestland and cropland groups, while annual precipitation determined the C source-sink relationship for the grassland group in which noticeable SOC sink strength was attributed to the conversion from cropped area to grass cover. Canonical correlation analysis also showed that grassland ecosystems are more complicated than others in the ecoregion, which may be difficult to identify on a field scale. Current model simulations need further adjustments to the model input variables for the grass cover-dominated ecosystems in the ecoregion.

Citation: Tan, Z., S. Liu, C. A. Johnston, J. Liu, and L. L. Tieszen (2006), Analysis of ecosystem controls on soil carbon source-sink relationships in the northwest Great Plains, *Global Biogeochem. Cycles*, *20*, GB4012, doi:10.1029/2005GB002610.

1. Introduction

[2] The spatial variability in ecosystem attributes present at landscape scales allows investigators to explore correlations between controls and response variables across broad regions and thereby to predict the response of ecosystems to global climate change. *Parton et al.* [1987] analyzed the factors controlling soil organic matter (SOM) levels in the Great Plains grasslands and found that the steady-state SOM level is a function of temperature, moisture, soil texture, and plant lignin. Their research laid a foundation for the development of the CENTURY SOM model [*Parton et al.*, 1993], which is widely used to estimate soil organic carbon (SOC) stocks on various spatial scales. However, *Burke et al.* [1997] pointed out that many of possible control factors covary with one another, and only some of the important factors actually exist in regional databases. Thus the true proximal controls may be difficult to identify on regional spatial scales. The simulation results of *Burke et al.* [1997] and the field observations of *Lauenroth and Sala* [1992] questioned the applicability of space-for-time substitutions when dealing with ecosystem function. On the basis of the investigation of spatial and temporal variations in net primary production and nitrogen (N) mineralization in the Great Plains grasslands, *Burke et al.* [1997] speculated that the structure of the systems may provide important constraints on their temporal variability that are not evident in an analysis of spatial variability, and that models describing spatial variability may not be appropriate to characterize temporal variability.

[3] *Tan et al.* [2005] reported that the SOC budgets and change rates over time in the northwest Great Plains were significantly correlated not only to the areal ratio of grassland to cropland, but also to the baseline SOC stock level. Moreover, the relationships of total SOC stock to each SOC fraction pool on the spatial scale differ significantly from those on the temporal scale. These results suggest that the baseline SOM level and fraction pools influence soil C trends.

[4] It is recognized that land use and land cover change (LUCC) plays an essential role in terrestrial C cycles [*Houghton et al.*, 1997, 1999] and global climate change [*Dale*, 1997; *Loveland et al.*, 2002]. First, these changes potentially transform the structure and function of ecosys-

[1]South Dakota Center for Biocomplexity Studies, Brookings, South Dakota, USA.
[2]Now at SAIC, USGS Center for Earth Resources Observation and Science (EROS), Sioux Falls, South Dakota, USA.
[3]SAIC, USGS Center for Earth Resources Observation and Science (EROS), Sioux Falls, South Dakota, USA.
[4]USGS Center for Earth Resources Observation and Science (EROS), Sioux Falls, South Dakota, USA.

0886-6236/06/2005GB002610$12.00

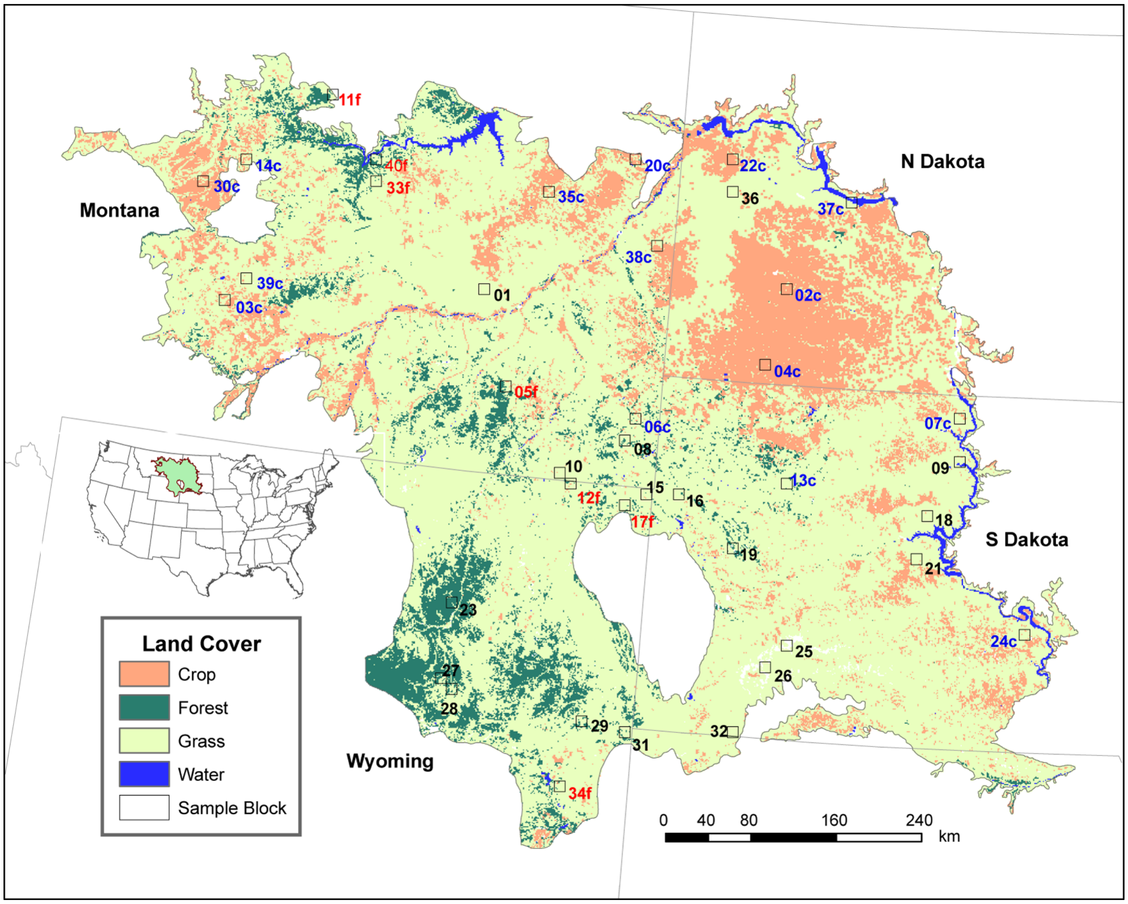

Figure 1. The northwest Great Plains ecoregion and sample block locations (integer numbers are the sample block IDs). (The IDs of sample blocks with "c" in blue refer to the cropland group, those with "f" in red refer to the forestland group, and the others are the grassland group.)

tems and alter the biogeochemical processes [*Post and Kwon*, 2000; *Reiners et al.*, 2002; *Liu et al.*, 2004b]. Second, LUCC has had a much greater impact on ecological variables than have climate changes [*Dale*, 1997]. Therefore driving forces associated with historical LUCC are required to properly evaluate terrestrial C dynamics, even though this kind of information is difficult to obtain over large areas.

[5] Despite the importance of LUCC in terrestrial C cycles, the quantitative relationships between terrestrial C budgets and LUCC trends are still poorly understood owing to the limited capability of current process-based models used to integrate spatial data with historical LUCC information and minimize simulation uncertainties for large areas [*Liu et al.*, 2004b]. Moreover, the relationships are not clear between the factors governing spatial variation and the controls responsible for temporal trends of terrestrial C fluxes. Current approaches, either direct measurements or process-based models, deal with individual ecosystem sectors separately. Thus they are incapable of including the interactions among individual ecosystems into an analysis of ecosystem controls for a whole ecoregion. The integration of historical LUCC information into model simulations contributes to an understanding of the actual controls on terrestrial C changes and the feedback on climate, since LUCC patterns and trends result from interactions among economic, environmental, social, political, and technological forces on local to global scales [*Loveland et al.*, 2002]. This study evaluated controls governing SOC dynamics over the period from 1973 to 2000 for major ecosystems in the northwest Great Plains of the United States.

2. Materials and Methods

2.1. Study Area

[6] The northwest Great Plains or Ecoregion 43 includes the parts of western South Dakota, southwestern North Dakota, southeastern Montana, and northeastern Wyoming, and covers an area of 338,718 km^2 (Figure 1). The average annual precipitation from 1973 to 2000 was 399 mm and the average annual temperature was 7.2°C. The annual maximum and minimum temperatures were 0.74 and 0.59°C higher between 1986 and 2000 than between 1973 and 1985, respectively. Land cover is dominated by mixed grasses, but agriculture is the primary land use transforming the grassland-dominated ecosystems. From 1973 through 2000, grass cover, cropped area, and forestry accounted for 75%, 17%, and 3%, respectively, of the total ecoregion area (Table 1). The cumulative change in land cover during the

Table 1. Summary of Preliminary Results for Basic Variables in the Northwest Great Plains

Group of Blocks	Number of Blocks	Mean/Stder[a]	Crop,[b] %	Forest,[b] %	Grass,[b] %	Precip., mm yr^{-1}	Temperature, °C yr^{-1} Max	Min	Mean	Baseline in 1973, g m^{-2} SOC	SON	Change,[c] % 1973–2000 SOC
Cropland	15	Mean	38.6	0.9	59.3	411	13.3	0.0	6.5	5823	481	−0.5
		Stder	6.1	0.4	5.9	11	0.3	0.1	0.2	475	40	1.7
Forestland	7	Mean	8.6	12.4	77.6	364	14.6	0.2	7.3	4854	395	1.8
		Stder	2.5	0.9	2.3	12	0.4	0.4	0.3	536	45	1.7
Grassland	18	Mean	2.8	0.5	94.4	404	15.1	0.0	7.7	3942	312	10.8
		Stder	0.8	0.3	0.7	9	0.3	0.2	0.2	199	16	1.4
Ecoregion	40	Mean	16.9	2.7	75.5	399	14.4	0.0	7.2	4807	390	5.3
		Stder	0.2	0.0	0.2	6	0.2	0.1	0.2	250	21	1.3

[a]Standard error.
[b]The percentage averaged from 1973 through 2000.
[c]Based on the baseline SOC stock in 1973.

period affected about 10% of the land area, but most of these changes were directly related to conversions between cropland and grassland, especially between 1986 and 2000 as a result of the Conservation Reserve Program (CRP) [*Tan et al.*, 2005].

2.2. Model Initialization and Simulation

2.2.1. General Ensemble Biogeochemical Modeling System

[7] The General Ensemble Biogeochemical Modeling System (GEMS) [*Liu et al.*, 2004a, 2004b] was used to simulate soil C dynamics in this study.

[8] GEMS is designed for regional-scale C estimation by integrating spatially explicit time-series LUCC change data into its simulations. As described by *Liu et al.* [2004a, 2004b], GEMS consists of three major components: single or multiple encapsulated ecosystem biogeochemical models, an automated stochastic parameterization system (AMPS), and an input/output processor (IOP). AMPS includes two major interdependent parts: the data search and retrieval algorithms and the data processing mechanisms. The first part searches for and retrieves relevant information from various databases according to the keys provided by a joint frequency distribution (JFD) table [*Reiners et al.*, 2002; *Liu et al.*, 2004b]. The data processing mechanisms downscale the aggregated information at the map-unit level to the field scale using a Monte Carlo approach. Once the data are assimilated, they are injected into the modeling processes through the IOP which updates the default input files with the assimilated data. Values of selected output variables are also written by the IOP to a set of output files after each model execution. The JFD grids are first created from soil maps, a time series of land cover images, and climate themes at a cell size of 60 m × 60 m.

[9] The spatial simulation unit of GEMS is a JFD case. A JFD case contains single or multiple, homogeneous, connected or isolated land pixels that represent a unique combination of values from the environmental Geographic Information System (GIS) layers used in an overlay operation [*Liu et al.*, 2004a]. These GIS layers include five dated land cover themes, soil and climate coverages, nitrogen (N) deposition map, and administrative districts. GEMS automates the processes of downscaling forest ages from the USDA Forest Inventory and Analysis data (FIA), crop compositions from the Agricultural Census, grass cover distribution and temporal changes from the remotely sensed imagery interpretation, and soil properties using a set of Monte Carlo processes. An ensemble simulation technique is employed to incorporate the uncertainties from the Monte Carlo processes.

[10] The data for model inputs primarily consisted of climatic regimes, LUCC, and soil inventory. The LUCC data were provided by the research team of the USGS Land Cover Trends Project [*Loveland et al.*, 2002]. Climatic data included annual precipitation and maximum and minimum temperature records from 1973 through 2000. They were converted to 30-m pixel GIS coverages from the CRU TS 2.0 data sets (available at http://www.cru.uea.ac.uk/cru/data/hrg.htm). Soil characteristics within each sample block were adapted from the U.S. State Soil Geographic Data Base (STATSGO) [*USDA-NRCS*, 1994] for initializing the soil components of GEMS.

2.2.2. Model Initialization

2.2.2.1. Crop Cultivation and Management

[11] In order to create JFDs for cropland, the Monte Carlo method was used to determine crop composition, on the basis of state-level averages of agricultural census data, because no detailed records exist for when and where crop species are planted in an ecoregion. Crop composition was assigned every five years following census data. As initialized for forest age [*Liu et al.*, 2004b], a crop species with a higher area percentage in the census data is more likely to be selected for simulation in the model. Between census times, crop rotation was specified on the basis of a transition probability table derived from census data. For example, if a JFD is assigned with spring wheat in the first year in one possible rotation, a 50% probability for spring wheat would follow in the next year, a 30% probability would be corn, and a 20% probability for other crops. We did not include the tillage management due to the difficulty in specifying the spatial distribution of county level data.

2.2.2.2. Grassland and Forestland

[12] Conversion between cropland and grassland has been a major land surface disturbance to grassland in the Ecoregion. Especially, the Conservation Reserve Program (CRP) since 1986 has stimulated conversion of cropped areas to grasslands, which may substantially shape the SOC dynamics. Fortunately, such conversions are detectable and accountable with time series remotely sensed images. For

other management practices such as grazing and fertilization, an average level for each was randomly assigned due to the lack of spatially explicit data on an ecoregion scale. Forestland was a small proportion of the total land area (about 2.7%); therefore the approach proposed by *Liu et al.* [2004a] was used to initialize forest age and biomass for model simulations.

2.2.2.3. Soil Characteristic Extraction and SOC Pool Partitioning

[13] Any JFD case uses soil information from a STATSGO map unit that contains different soil characteristics. A Monte Carlo method was used to assign each JFD a set of specific soil property values such as layer depth, soil organic matter content, soil water holding capacity, and clay and sand percentages. Soils with high SOC contents are usually fertile and likely lead to high net primary productivity (NPP), which may cause high C release. Therefore it is important to properly partition SOC storage into different pools at the beginning of each simulation. Previously, the initialization of each SOC pool required a spinup simulation over a long-term to find a soil C equilibrium for undisturbed vegetation. The reconstructed disturbance history was then used to get a close estimate of the SOC pools. This requires informative historical data which are usually not available for a large area. By testing the model structure, we found that the slow SOC pool is almost linearly proportional to the NPP level. This result was noticed by *Paustian et al.* [1995] who observed a linear relationship between SOC level and C input from plant residues. On the basis of these findings, we used a retrospective SOC initialization algorithm to define the slow SOC pool on the basis of the NPP for each land cover type and soil inventory data. The difference between the total SOC and the slow pool was then used to initialize the passive SOC pool. The active SOC pool was set at about 2% of the total SOC storage.

2.2.3. Ensemble Simulations

[14] GEMS generates site-level inputs with a Monte Carlo approach from regional data sets. Any single simulation of a JFD case is unique combination of randomly picked forest age, crop species, and soil properties from regional-level data sets, so that the output of a single simulation run of a JFD might be biased. Therefore ensemble simulations of each JFD were executed to incorporate the variability of inputs and to average uncertainties of simulation results. In general, averages of ensemble simulations become more stable when increasing the run number. We made 20 repeat runs for each JFD case in this study, which reduced the relative error to about 2%. The averaged JFD output from the 20 runs was then aggregated on sample block scale, and the simulation uncertainty was evaluated on both sample block and the ecoregion scales.

2.3. Sample Blocks

[15] Forty sample blocks of 10 km × 10 km, randomly selected within the ecoregion by *Loveland et al.* [2002] (Figure 1), were used to identify changes with 1% precision at an 85% confidence level. The changes were detected on the basis of five calendar years (1973, 1980, 1986, 1992, and 2000) of Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM) image data. They were analyzed at a cell size of 60 m × 60 m for MSS images and 30 m × 30 m for TM images.

2.4. Data Analyses

[16] Our preliminary study showed that the SOC dynamics were differentiated by the land cover composition within each sample block. Therefore the summarization and analyses of model outputs were conducted with sample block groups defined sequentially using the average land cover proportions between 1973 and 2000 as follows: forestland group if the forested area in a sample block was equal or greater than 10%, regardless of the other land cover proportions; then cropland group if the areal proportion of the cropped area in the sample blocks was greater than 10%; finally, remaining blocks were assigned to the grassland group. Seven of the sample blocks fell into the forestland group, fifteen into the cropland group, and eighteen into the grassland group (Figure 1 and Table 1). We used sample block groups as data aggregation units, and analyzed them for a time series from 1973 through 2000. The C stocks in g C m^{-2} were defined as the total SOC pools in the 0 to 20 cm depth of soil. The SOC change was determined as the difference in SOC stocks between 1973 and 2000.

[17] Multiple regression analysis was conducted using SAS [*SAS Institute*, 2003] to examine the dependency of the SOC change between 1973 and 2000 upon selected explanatory variables (Table 1). Canonical correlation analysis (CCA) was run to identify the dominant associations between sets of control and response variables, and to determine the extent to which the variation of response variables could be attributed to controls [*Töth et al.*, 1995; *Dieleman et al.*, 2000].

3. Results and Discussion

3.1. Soil Organic C Source-Sink Relationships Between 1973 and 2000

[18] The results presented in Table 1 show large differences in SOC changes among the three sample block groups. On average, there was 0.5% decrease in SOC stock for the cropland group from 1973 to 2000, and 1.8% and 10.8% increases for the forestland group and grassland group, respectively. However, these changes varied significantly among sample blocks within each land cover group (Figure 2). Carbon sources were mainly associated with the sample blocks where cropped area proportion was dominant, and C sinks typically occurred in the sample blocks where grass cover was prevalent. As shown in Figure 2, larger sinks were associated with sample blocks where the average baseline SOC levels were lower; whereas larger sources were linked to the sample blocks with higher baseline SOC levels. Carbon sinks or sources were dominantly attributed to changes in the slow and labile C pools, especially for the sample blocks with C sources.

3.2. Forces Driving SOC Dynamics

3.2.1. Baseline SOC Stock

[19] As illustrated in Figure 3, the strength of SOC sinks or sources was proportional to the baseline SOC stock level. The rate of SOC augmentation tended to decline with an

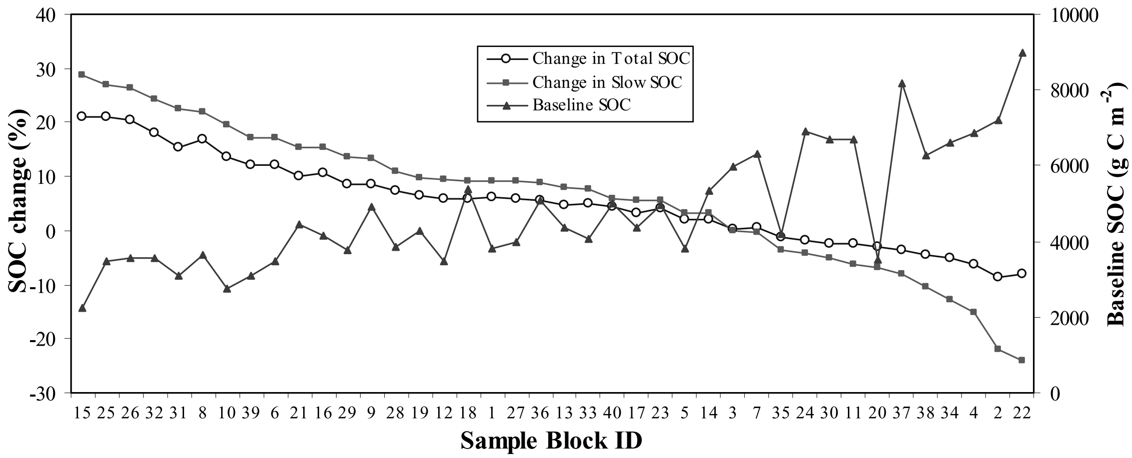

Figure 2. Changes in total soil organic C (SOC) and slow C pools between 1973 and 2000 and their relations to the baseline SOC stock in the ecoregion (presented in the descending order of change rates). The changes (%) in the SOC and slow C pools between 1973 and 2000 were based on the baseline SOC stock for each sample block in 1973.

increase in high baseline SOC levels. As observed by *Conant et al.* [2003] and *Tan and Lal* [2005], higher C sequestration rates are associated with soils having most recently undergone conservation management, but the rates tend to decrease with time. These results suggest that either C sinks or sources for individual ecosystems are to a great extent dependent on the baseline SOC stock level, which is also seen from the changes in the slow fraction pool (Figure 2).

3.2.2. Land Cover Composition

[20] Cropped area proportion in five sample blocks (02, 04, 20, 22, and 30) of the 15 blocks in the cropland group was greater than 50% of the block area. The average grass cover proportion across a sample block in the grassland group accounted for 94% of the block area. The change in all land cover types within the ecoregion took place at an average annual rate of 0.35% between 1973 and 2000 [*Tan et al.*, 2005] and there were significant differences among sample blocks. For example, 1.34% and 1.02% annual changes occurred within sample blocks 04 and 20, respectively. As mentioned above, the principal LUCC in the ecoregion was the conversion between cropland and grassland. Therefore we used the change in land cover composition ratio of both grass and crop covers to represent land cover dynamics. On the sample block scale, we observed that the SOC changes between 1973 and 2000 were logarithmically proportional to the areal ratio of grass cover to cropland (Figure 4); the C sources became stronger whenever the ratio was below 2, as proposed by *Tan et al.* [2005], supporting the hypothesis that the conversion from grassland to cropland likely accelerates SOC depletion.

3.2.3. Differences in SOC Dynamics Among Three Sample Block Groups

[21] The changes in SOC stocks over time were illustrated by land cover composition groups in Figure 5. The cropland group (with an average cropped area proportion of 39%) had an average baseline SOC stock of 5823 g C m^{-2} and was a small C source (3.9 g C m^{-2}) over the 28-year span.

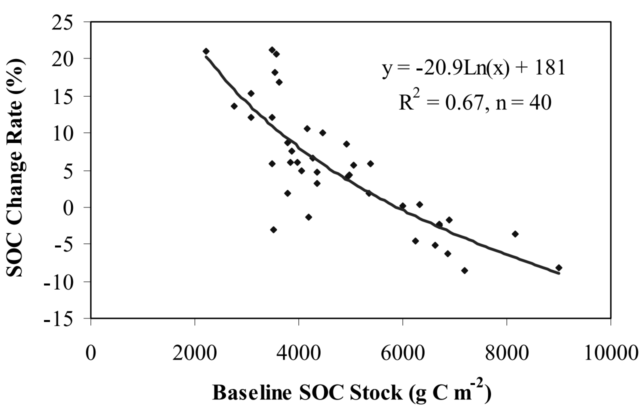

Figure 3. SOC change rates from 1993 to 2000 in relation to the baseline SOC stock in 1973, illustrated by 40 sample blocks.

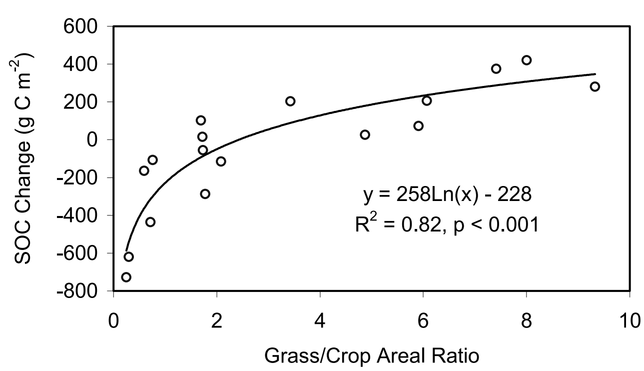

Figure 4. Effects of land cover composition on soil organic C sink-source relationships between 1973 and 2000 for sample blocks where cropped area percentage was >10% (positive SOC change represents a C sink).

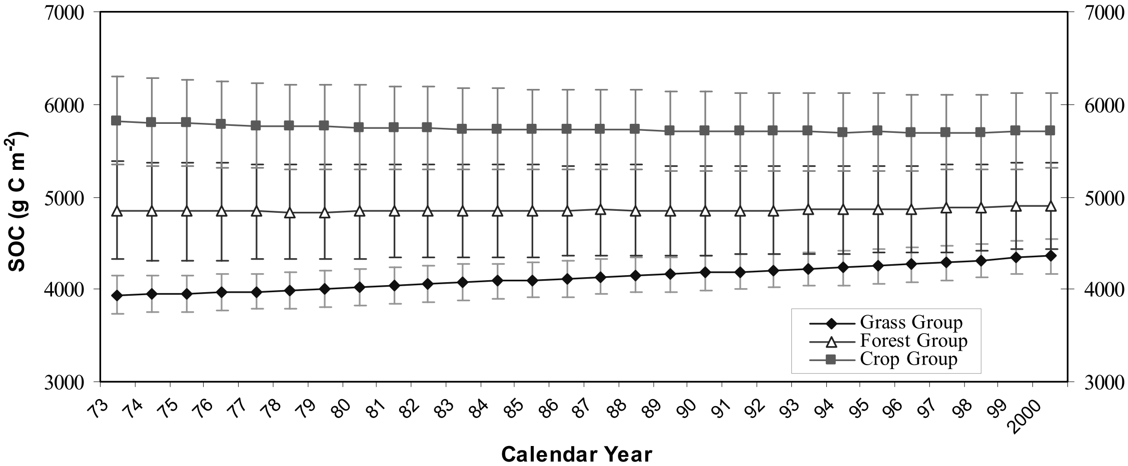

Figure 5. Temporal variation of soil organic C and uncertainty as related to land cover composition (sample block group) in the ecoregion (error bars are standard errors).

The grassland group had the largest C sink strength (14.9 g C m^{-2}) probably because of the low baseline SOC level (3942 g C m^{-2}) and other contributors (see discussion below). Both the baseline and the change rate of SOC stock showed significant differences among these three groups of sample blocks ($p < 0.05$). Note that the baseline SOC data from STATSGO was not based on land cover types; land cover distribution is usually oriented to human needs. Croplands are generally located on Mollisols and some Vertisols where the SOC contents are high and the relief is level or gentle sloping. Forestlands are mainly located on Aridisols and Alfisols. Grasslands are primarily located on Entisols and Inceptisols in southwestern part of the ecoregion, where there is less SOC storage owing to steep slopes and sparse vegetation cover in comparison with forested and cropped areas. Again, some portion (<10%) of cropped areas existed within the grassland group and some defined grass cover was converted from previous cropland. As indicated by Figure 1, croplands are sporadically distributed within grasslands. They have higher fertility and biophysical settings that favor crop production relative to neighboring parcels. These sparsely cropped parcels were also prone to be included in the CRP and converted to grassland, which would enhance C sequestration due to the reduced disturbance under grass cover. As a result of the conversions, the C sink strength was 35% higher between 1986 and 2000 than between 1972 and 1986 (and C source strength on croplands decreased from 7.0 g C $m^{-2}yr^{-1}$ to 1.0 g C $m^{-2}yr^{-1}$). This, along with the data in Figure 6, suggests that the higher C sink strength in the grassland

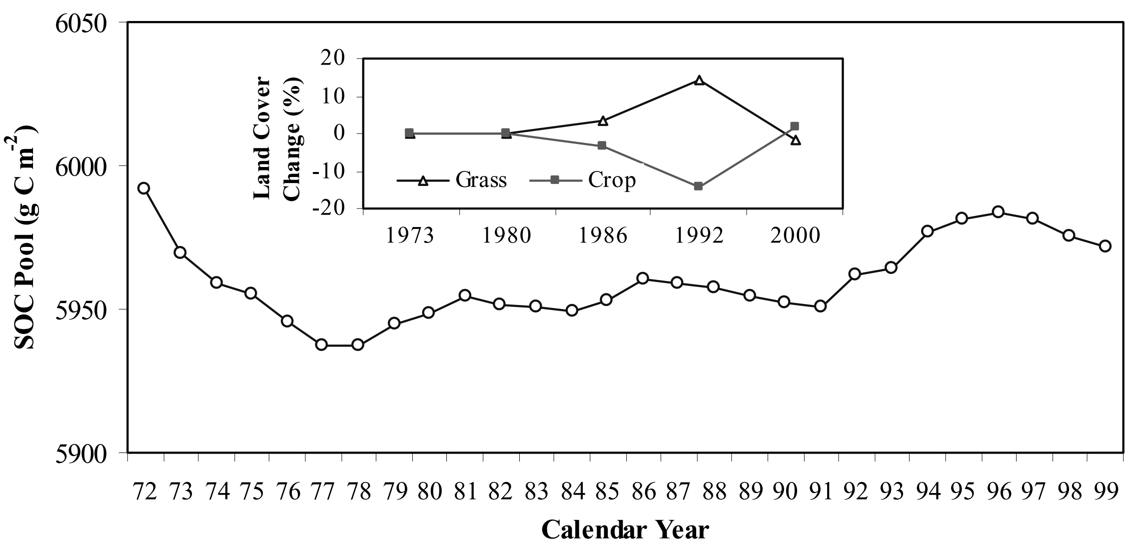

Figure 6. SOC pool trends over time as related to the conversion between grassland and cropland in sample block 3. Embedded graph indicates the land cover change detected from remotely sensed images of five dates.

Table 2. Standardized Canonical Coefficients (SCC) for the Controls on SOC Stock Change Between 1973 and 2000 in the Ecoregion[a]

Group of Blocks	Control	SCC	Standardized Variance of SOC Change Explained By Canonical Variables of Controls
Ecoregion	Grass	0.52	90%, $p < 0.0001$
	SON_{bs}	−0.29	90%, $p < 0.0001$
	SOC_{bs}	−0.27	90%, $p < 0.0001$
	PPT	0.23	90%, $p < 0.0001$
	T_{min}	0.07	90%, $p < 0.0001$
	T_{max}	−0.03	90%, $p < 0.0001$
Cropland	SOC_{bs}	−1.32	90%, $p = 0.0012$
	SON_{bs}	0.58	90%, $p = 0.0012$
	CGR	−0.40	90%, $p = 0.0012$
	PPT	0.34	90%, $p = 0.0012$
	T_{min}	−0.13	90%, $p = 0.0012$
	T_{max}	0.02	90%, $p = 0.0012$
Forestland	SON_{bs}	6.39	100%
	SOC_{bs}	−5.88	100%
	Forest	0.74	100%
	T_{max}	0.62	100%
	T_{min}	0.21	100%
	PPT	−0.05	100%
Grassland	PPT	0.81	73%, $p = 0.03$
	Crop	−0.59	73%, $p = 0.03$
	SON_{bs}	−0.46	73%, $p = 0.03$
	T_{min}	0.38	73%, $p = 0.03$
	Grass	0.13	73%, $p = 0.03$
	SOC_{bs}	−0.13	73%, $p = 0.03$
	T_{max}	−0.09	73%, $p = 0.03$

[a]Crop, cropland areal percentage; GCR, areal ratio of grassland over cropland; Grass, grassland areal percentage; Forest, forestland areal percentage; T_{max}, annual mean maximum temperature (°C); T_{min}, average annual minimum temperature (°C); PPT, annual precipitation (mm yr^{-1}); SCC, standardized canonical coefficient; SOC_{bs}, total soil organic carbon (g C m^{-2}); SON_{bs}, total soil organic nitrogen (g N m^{-2}).

group could be, to a great extent, attributed to the conversion from crop cultivation to grass cover.

3.3. Evaluation of Controls

3.3.1. Multiple Regression Analysis

[22] To test the hypothesis that changes in SOC stock between 1973 and 2000 were mainly a function of land cover composition, climate variables, and the baseline levels of SOC and SON stocks, we conducted a linear regression analysis. The difference in SOC stock can be well predicted using these variables in following equation:

$$\begin{aligned} \mathrm{SOCD} = & -537 - 0.06(\mathrm{SOC_{bs}}) - 0.72(\mathrm{SON_{bs}}) + 8.1(\mathrm{Grs}) \\ & + 1.9(\mathrm{PPT}) - 8.5(\mathrm{T_{max}}) + 31(\mathrm{T_{min}}) \\ & \left(p < 0.0001, R^2 = 0.90, n = 40\right), \end{aligned} \tag{1}$$

where SOCD is the difference in SOC stock between 1973 and 2000 (g C m^{-2}) estimated on the basis of all 40 sample blocks; SOC_{bs} and SON_{bs} are the baseline stocks of respective SOC and SON in 1973 (g C m^{-2}); Grs is the areal percentage of grassland (%) for the period from 1973 through 2000; PPT is average annual precipitation (mm); T_{max} and T_{min} are the average maximum and minimum monthly temperature (°C), respectively.

[23] Equation (1) indicates that more precipitation, higher minimum temperature, and larger proportion of grass cover favored the SOC accumulation in the ecoregion, while the rate was negatively correlated to the baseline soil organic matter stock and the maximum temperature. The power analysis shows that, of these variables, the grass cover proportion, baseline SOC stock and annual precipitation played a dominant role in the SOC budget. The influences of the other factors are very limited. The power index magnitude order is: SOC_{bs} (1.0) = Grs (1.0) ≈ Rain (0.99) ≫ SON_{bs} (0.25) > T_{min} (0.12) > T_{max} (0.05).

3.3.2. Canonical Correlation Analysis

[24] Although SOC changes could be well predicted by the variables listed in equation (1), we cannot judge the importance of each variable to the response. Therefore a canonical correlation analysis was conducted for this purpose. The magnitude and sign of a standardized canonical coefficient (SCC) indicates the contribution of a control variable to its canonical variate [*Dieleman et al.*, 2000; *Tan et al.*, 2004]. The first canonical variate of control variables for each sample block group was expressed by its original control variables as presented in Table 2. We used baseline SOC and SON stocks, grass cover proportion, cropped areal proportion, annual precipitation, and minimum and maximum temperature regimes as the control variables to explain the variance of the SOC change from 1973 to 2000. We found that these control variables explained 90% of such variance at $p < 0.0001$ level. Of all selected controls, the areal proportion of grass cover played an important role in determining C sources or sinks overall, but the C sink strength tended to become weaker with an increase in the baseline SOC and SON levels. In other words, C sources more likely occur in soils with higher baseline SOC and SON contents. This conclusion is consistent with the conclusion drawn by other researchers [*Jastrow and Miller*, 1997; *Six et al.*, 2002; *West and Post*, 2002; *Tan and Lal*, 2005; *Tan et al.*, 2005].

[25] Further analyses showed that the controls on and their importance to the C source-sink relationship varied with different sample block groups (Table 2). For the cropland group, the SOC budget was mainly related to the baseline SOC contents, followed by the baseline SON, the ratio of cropped area to grass cover, and annual precipitation in descending order of importance. And the C sink strength likely increased with baseline SON and annual precipitation, whereas the C sources tended to occur in soils with high baseline SOC because high C soils are likely choose for cropland. For the forestland group, both baseline SOC and SON predominate the soil C source-sink relationship even though higher forest cover proportion and temperature regimes would favor C accumulation in soils with either higher baseline SON levels or lower baseline SOC stocks.

[26] For the grassland group, the annual precipitation was the most important force driving the SOC budget. A high annual precipitation along with a high minimum temperature regime would enhance soil C sequestration, but the sink strength could be weakened in either cropped areas or in soils with high baseline SON contents. The combination of all variables explained only 73% of the standardized variance of SOC change (Table 2). Note that similar controls on SOC change (shown by the canonical correlation analysis listed in Table 2 for these three sample block groups) do not

necessarily mean that management measures have little impact on C dynamics. Each of these sample block groups was a mosaic of different land cover types in which the C sequestration enhanced by one measure could be offset by the C depletion induced by another.

[27] It is understandable that the cropping-dominated ecosystem is the most extensively managed since chemicals and irrigation would be preferentially used to control biomass and ameliorate climatic effects. For example, the irrigated cropland in 1998 was 29% of all harvested cropland area for the ecoregion and >50% across North Dakota state (http://www.ers.usda.gov/Data/WesternIrrigation/). The forest ecosystem generally experiences the least human-induced disturbance and its SOC dynamics would be more likely driven by nonmanagement factors. In comparison to the above two ecosystems, the SOM retention on grassland would be much more influenced by complex relationships among climate, soil microbes, grass species, and livestock; additionally, the aboveground biomass and the belowground organic matter accumulation are more dependent upon climate regimes. Meanwhile, the SOC depletion induced by erosion on grassland may be more serious than on forestland and cropland owing to the differences in the combination of surface cover and relief. For example, grasslands are usually on the steeper slopes in comparison to croplands. A considerable unexplained variance for the grassland block group (27%) may be attributed to variables which were not included in the model inputs. Moreover, the uncertainties of the driving forces associated with the native grass distribution and varying grazing intensities make the grassland ecosystems more complicated than both the forest and cropping ecosystems in this ecoregion. In other words, the conventional paradigm that grassland ecosystems are simpler than either forestland or cropland ecosystems may only be true on the field scale, not on landscape or regional scales.

[28] Our analysis showed that the baseline SOC level is a good predictor of soil C dynamics. Probably, SOM storage collectively results from long-term interactions of climate variables (precipitation and temperature), vegetation (lignin), relief (drainage condition), and soil parent materials (geologic origin and texture) [*Jenny*, 1980; *Parton et al.*, 1987; *Schimel et al.*, 1994; *Tan et al.*, 2003]. The data in Table 2 suggest that SOC change depends more upon the baseline SOC level than other factors for the described historical land use scenarios, which is particularly explicit for the forestland group. However, this does not mean that LUCC or conversions between land covers are not critical to soil C dynamics. Instead, it underscores the importance of the proportion of changes in the land cover within the entire region, even though the changes in C stocks are usually stimulated by land cover change, and the direction of conversion between grassland and cropland likely determines whether a site is a sink or a source [*Tan et al.*, 2005]. Applying a Monte Carlo approach, randomly selected reference SOC stocks, and management factors to the IPCC model, *Ogle et al.* [2003] reported that losses of SOC from 1982 to 1997 on US agricultural lands were mainly from managed organic soils and gains came from managed mineral soils. Their results, along with those from the cropland group in this study (Table 1), suggest that soils with higher baseline C stocks tend to become larger C sources under cultivation, and soils with lower baseline C levels likely become C sinks with conservation and improved management practices. A similar conclusion was also drawn by *Tan and Lal* [2005] on the basis of field observations in the north-central U.S.

[29] *Cannell and Thornley* [1998] and *Reich et al.* [2001] reported that the N-poor grasslands tend to undergo a larger long-term response to elevated CO_2 owing to a slower N accumulation coupled with lesser N leaching, decreased gaseous N loss, and increased N_2 fixation. Additionally, if more explanatory variables were involved in model inputs, the magnitudes and signs of SCC values for respective variables could be altered. These may be the reasons why the baseline SON level showed a negative relation to the change rate of total SOC stock on the grassland group for the modeled time span.

[30] The role of climatic variables in SOC dynamics for large areas is widely recognized. Generally, SOC stocks increase with precipitation and decrease with temperature [*Jenny*, 1980; *Burke et al.*, 1989]. In the temperate forestland of Minnesota, Wisconsin, and Michigan, SOC stocks increase with mean annual precipitation [*Grigal and Ohmann*, 1992]. Across the Great Plains Grassland, SOC stocks are positively correlated with annual precipitation and negatively correlated with annual mean temperature [*Burke et al.*, 1989]. However, the climatic impacts seem to be complicated for the study area because they may have been reflected in the baseline SOM levels. Meanwhile, the interpretation of temperature effects on SOC change may need to consider the influence of temperature regimes on vegetation net primary production. In this ecoregion, mean annual temperature is only 7.1°C, ranging from 0.0° to 14.4°C; the mean annual maximum temperature is lower than the optimal temperature range for plant growth. An elevation in maximum temperature in this ecoregion could enhance biomass production more than consumption by respiration. An increase in minimum temperature would lead to an increase in SOC stock for the grassland group, since higher minimum temperature likely leads to a longer growing season and higher biomass. Overall, our results agree with the conclusion that generally weak relationships between climatic variables and SOC stocks on a regional scale make it difficult to predict changes in SOC stock as a function of projected climate change only [*Kern et al.*, 1998].

3.4. Model Validation

[31] GEMS simulates C biogeochemical processes on both spatial and temporal scales, and is based on applying the well-established CENTURY SOM model in space using the JFD of major driving variables of the C cycle [*Liu et al.*, 2004a, 2004b]. The inputs for and outputs from GEMS simulations are more representative of the spatial and temporal heterogeneity of the driving variables than those based on the wall-to-wall simulation that ignores the spatial explicitness and covariance of these variables for large areas [*Tan et al.*, 2005]. GEMS simulations were processed for each case, consisting of a randomly chosen combination of

land cover and soil taxon with respective inputs retrieved from JFD files; each case was run 20 times to create outputs weighted by area proportion of cases with standard deviations. Uncertainties with the GEMS outputs also indicate that 40 sample blocks was enough to capture the general spatial and temporal variability of C fluxes and pools across the ecoregion. Conventional validation of GEMS is not possible due to the lack of dynamic regional-scale SOC databases. It is also impossible to validate GEMS simulation results using limited point field measurements [*Tan et al.*, 2005]. However, the grain yields of major crops estimated from GEMS were quite consistent with the statewide mean values provided by the USDA National Agricultural Statistical Service [*Tan et al.*, 2005]. Moreover, a good match was observed between the historical SOC trends and the temporal pattern of land cover change (Figure 6). These results suggest the stability and robustness of the GEMS-CENTURY model.

[32] **Acknowledgments.** This research was part of the Earth Surface Dynamics Program and the Geographic Analysis and Monitoring Program of the USGS, and also supported by a NASA grant (LUCC99-0022-0035). The authors appreciate Zhengpeng Li's assistance with the model simulations and thank all reviewers. Part of the work (Z. Tan, S. Liu, and J. Liu) was performed under U.S. Geological Survey contract 03CRCN0001.

References

Burke, I. C., C. M. Yonker, W. J. Parton, C. V. Cole, K. Flach, and D. S. Schimel (1989), Texture, climate, and cultivation effects on soil organic matter content in US grassland soils, *Soil Sci. Soc. Am. J.*, *53*, 800–805.

Burke, I. C., W. K. Lauenroth, and W. J. Parton (1997), Regional and temporal variation in net primary production and nitrogen mineralization in grasslands, *Ecology*, *78*, 1330–1340.

Cannell, M. G. R., and J. H. M. Thornley (1998), N-poor ecosystems may respond more to elevated [CO_2] than N-rich ones in the long term: A model analysis of grassland, *Global Change Biol.*, *4*, 431–442.

Conant, R. T., J. Six, and K. Paustian (2003), Land use effects on soil carbon fractions in the southeastern United States: I. Management-intensive versus extensive grazing, *Biol. Fertil. Soils*, *38*, 386–392.

Dale, V. H. (1997), The relationship between land-use change and climate change, *Ecol. Appl.*, *2*, 753–769.

Dieleman, J. A., D. A. Mortensen, D. D. Buhler, C. A. Cambardella, and T. B. Moorman (2000), Identifying associations among site properties and weed species abundance: I. Multivariate analysis, *Weed Sci.*, *48*, 567–575.

Grigal, D. G., and L. F. Ohmann (1992), Carbon storage in upland forests of the Lake States, *Soil Sci. Soc. Am. J.*, *56*, 935–943.

Houghton, J. T., L. G. Meira Filho, B. Lim, K. Tre'anton, I. Mamaty, Y. Bonduki, D. J. Griggs, and B. A. Callander (Eds.) (1997), *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories*, 3 vols., Hadley Cent. Meteorol. Off., Bracknell, UK.

Houghton, R. A., J. L. Hackler, and K. T. Lawrence (1999), The U. S. carbon budget: Contributions from land-use change, *Science*, *285*, 574–578.

Jastrow, J. D., and R. M. Miller (1997), Soil aggregate stabilization and carbon sequestration: Feedbacks through organomineral associations, in *Soil Processes and the Carbon Cycle*, edited by R. Lal et al., pp. 207–223, CRC Press, Boca Raton, Fla.

Jenny, H. (Ed.) (1980), *The Soil Resource: Origin and Behavior*, 377 pp., Springer, New York.

Kern, J. S., D. P. Turner, and R. F. Dodson (1998), Spatial patterns in soil organic carbon pool size in the northwestern United States, in *Soil Processes and the Carbon Cycle*, edited by R. Lal et al., pp. 29–43, CRC Press, Boca Raton, Fla.

Lauenroth, W. K., and O. E. Sala (1992), Long-term forage production of North American shortgrass steppe, *Ecol. Appl.*, *2*, 397–403.

Liu, S., M. Kaire, E. Wood, O. Dialloc, and L. L. Tieszen (2004a), Impacts of land use and climate change on carbon dynamics in south-central Senegal, *J. Arid Environ.*, *59*, 583–604.

Liu, S., T. R. Loveland, and T. R. Kurtz (2004b), Contemporary carbon dynamics in terrestrial ecosystems in the southeastern plains of the United States, *Environ. Manage.*, *33*, 442–456.

Loveland, T. R., T. L. Sohl, S. V. Stehman, A. L. Gallant, K. L. Sayler, and D. E. Napton (2002), A strategy for estimating the rates of recent United States land-cover changes, *Photogramm. Eng. Remote Sens.*, *68*, 1091–1099.

Ogle, S. M., F. J. Breidt, M. D. Eve, and K. Paustian (2003), Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997, *Global Change Biol.*, *9*, 21–1542.

Parton, W. J., D. S. Schimel, C. V. Cole, and D. S. Ojima (1987), Analysis of factors controlling soil organic matter levels in Great Plains Grasslands, *Soil Sci. Soc. Am. J.*, *51*, 1173–1179.

Parton, W. J., et al. (1993), Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, *Global Biogeochem. Cycles*, *7*, 785–809.

Paustian, K., G. P. Robertson, and E. T. Elliott (1995), Management impacts on carbon storage and gas fluxes (CO_2, CH_4) in mid-latitude cropland and grassland ecosystems, in *Soil Management and Greenhouse Effect*, edited by R. Lal et al., pp. 69–84, CRC Press, Boca Raton, Fla.

Post, W. M., and K. C. Kwon (2000), Soil carbon sequestration and land-use change: Processes and potential, *Global Change Biol.*, *6*, 317–327.

Reich, P. B., et al. (2001), Do species and functional groups differ in acquisition and use of C, N and water under varying atmospheric CO_2 and N availability regimes? A field test with 16 grassland species, *New Phytol.*, *150*, 435–448, doi:10.1046/j.1469-8137.2001.00114.x.

Reiners, W. A., S. Liu, K. G. Gerow, M. Keller, and D. S. Schimel (2002), Historical and future land use effects on trace gas emissions using an ensemble modeling approach: Costa Rica's Caribbean Lowlands as an example, *Global Biogeochem. Cycles*, *16*(4), 1068, doi:10.1029/2001GB001437.

SAS Institute (2003), *SAS User's Guide, Release 9.1*, Cary, N. C.

Schimel, D. S., B. H. Braswell, E. A. Holland, R. McKeown, D. S. Ojima, T. H. Painter, W. J. Parton, and A. R. Townsend (1994), Climatic, edaphic, and biotic controls over storage and turnover of carbon in soils, *Global Biogeochem. Cycles*, *8*, 279–293.

Six, J., R. T. Conant, E. A. Paul, and K. Paustian (2002), Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils, *Plant Soil*, *241*, 155–176.

Tan, Z., and R. Lal (2005), Carbon sequestration potential estimates with changes in land use and tillage practice in Ohio, USA, *Agric. Ecosyst. Environ.*, *111*, 140–152, doi:10.1016/j.agee.2005.05.012.

Tan, Z., R. Lal, N. E. Smeck, F. G. Calhoun, R. M. Gehring, and B. Parkinson (2003), Identifying associations among soil and site variables using canonical correlation analysis, *Soil Sci.*, *168*, 376–382.

Tan, Z., R. Lal, N. E. Smeck, F. G. Calhoun, B. Parkinson, and R. M. Gehring (2004), Relationships between soil organic carbon pool and site variables in Ohio, *Geoderma*, *121*, 187–195.

Tan, Z., S. Liu, C. A. Johnston, T. R. Loveland, L. L. Tieszen, J. Liu, and R. Kurtz (2005), Soil organic carbon dynamics as related to land use history in the northwestern Great Plains, *Global Biogeochem. Cycles*, *19*, GB3011, doi:10.1029/2005GB002536.

Töth, T., S. Matsumoto, R. Mao, and Y. Yin (1995), Precision of predicting soil salinity based on vegetation categories of abandoned lands, *Soil Sci.*, *160*, 218–231.

USDA-NRCS (1994), State Soil Geographic (STATSGO) Data Base: Data use information, *Misc. Publ. 1492*, 110 pp., Washington, D. C.

West, T. O., and W. M. Post (2002), Soil organic carbon sequestration rates by tillage and crop rotation: A global data analysis, *Soil Sci. Soc. Am. J.*, *66*, 1930–1946.

C. A. Johnston, South Dakota Center for Biocomplexity Studies, Brookings, SD 57007, USA.

J. Liu, S. Liu, and Z. Tan, SAIC, USGS Center for Earth Resources Observation and Science (EROS), Sioux Falls, SD 57198, USA. (ztan@usgs.gov)

L. L. Tieszen, USGS Center for Earth Resources Observation and Science (EROS), Sioux Falls, SD 57198, USA.